Coupled Global-Regional Data Assimilation Using Joint States

Istvan Szunyogh
Texas A&M University, Department of Atmospheric Sciences
3150 TAMU
College Station, TX, 77843-3150
phone: (979) 458-0553 fax: (979) 862-4466 email: szunyogh@tamu.edu

Award Number: N00014-12-1-0785
http://atmo.tamu.edu/profile/ISzunyogh

LONG-TERM GOALS

The main goal of this research project is to develop a data assimilation system to obtain a global atmospheric analysis for the U. S. Navy's Operational Global Atmospheric Prediction System (NOGAPS) model, as well as a set of limited area atmospheric analyses for multiple local domains for the Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) model by a single data assimilation process. We will achieve this goal by developing a novel data assimilation system based on analyzing the joint states of the global and the limited area models.

OBJECTIVES

The Fleet Numerical Meteorology and Oceanography Center (FNMOC) prepares both global and limited area weather analyses and forecasts. In fact, FNMOC prepares limited area model products for more regions (more than 60) than any other center in the world. In the current implementation of the model suite, the global model is started from analyses prepared by the Naval Research Laboratory Atmospheric Variational Data Assimilation--Accelerated Presenter (NAVDAS-AR) data assimilation system, which is based on a 4D-VAR data assimilation scheme, while the regional model is started from analyses provided by the Naval Research Laboratory Atmospheric Variational Data Assimilation (NAVDAS) system for the atmosphere and the Navy Coupled Ocean Data Assimilation (NCODA) for the ocean. Both NAVDAS and NCODA are 3D-VAR schemes. In this configuration, deterministic model information is propagated from the global model to the regional analysis through the lateral boundary conditions. Building on the results of our earlier research, we are developing a data assimilation algorithm, in which information flows in both directions between the global and the limited area data assimilation systems. We expect both the global and the limited area analyses to benefit from the coupled approach. In particular, we expect that in the coupled data assimilation system, the global analyses will benefit from the availability of the high-resolution limited area model information in regions where the presence of small scale atmospheric flow features (e.g., in a tropical cyclone or over complex terrain) severely restrict the representativeness of the observations at the scales resolved by the global model.

In addition, we hope that in the process of developing and testing the data assimilation system, we will gain new knowledge about the mechanisms by which mesoscale processes influence synoptic and

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **30 SEP 2013**	2. REPORT TYPE	3. DATES COVERED **00-00-2013 to 00-00-2013**
4. TITLE AND SUBTITLE **Coupled Global-Regional Data Assimilation Using Joint States**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Texas A&M University,Department of Atmospheric Sciences,3150 TAMU,College Station,TX,77843-3150**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**11**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

global scale predictability. Such new knowledge will help make strategic decisions about the development of the analysis and forecast systems of the future.

APPROACH

Our approach takes advantage of the results of our earlier theoretical and modeling efforts on coupling the global and the limited area data assimilation processes (Merkova et al. 2011; Holt 2011; Holt et al. 2013; Yoon et al. 2012) and the advances made by Dr. Craig Bishop and his NRL Monterey-based research group by incorporating ensemble-based flow-dependent information into NAVDAS. Dr. Bishop and his group have developed an ensemble-based, highly portable version of NAVDAS for NOGAPS and COAMPS. We work in close collaboration with Dr. Bishop's group. The relevance of our research is expected to highly benefit from using a state-of-the-art operational system that includes capabilities to assimilate satellite radiance observations and to perform normal mode initialization. In addition, using the NRL system is expected to greatly accelerate the transfer of the research results to NRL, Monterey, and eventually to FNMOC.

WORK COMPLETED

Since the legal process to gain access to the model and data assimilation codes from the Naval Research Laboratory, Monterey, CA, has taken longer than expected, our efforts have been primarily focused on further developing the theory and testing the resulting new ideas on the coupled global-limited-area data assimilation system described in Holt et al. 2013. In particular, we made progress in three areas

- Development of the theory of the joint-state approach for the situation in which multiple limited area domains can intersect.
- Development of a robust version of the ensemble-based Kalman filter for online quality control of the observations.
- Testing of the new ideas, including the robust version of the ensemble-based Kalman filter, by assimilating Tropical Cyclone observations by the global-limited-area data assimilation system.

In what follows, we summarize the main results for each research area.

RESULTS

Multiple Limited Area Domains

This part of the research has been carried at the University of Maryland (UMD) under a subcontract from Texas A&M University. The Principal Investigators at UMD are Profs. Edward Ott and Brian Hunt, while Matthew Kretschmer, a graduate student, carried out the idealized model calculations.

We consider the case where each limited area model (LAM) is producing an ensemble forecast, and performs DA locally, using an algorithm known as the Local Ensemble Transform Kalman Filter (LETKF, Hunt et al. 2007). The LETKF produces an analysis ensemble by finding an optimal linear combination of the background ensemble members. This information is given in terms of a weight vector **w** for the analysis, and a weight matrix **W** for the analysis ensemble. These two quantities are,

in general, slowly varying from grid point to grid point, and have the same dimensionality as the ensemble.

The starting point of our technique is the joint states analysis method of Yoon et al (2012), which simultaneously performs an analysis using the LETKF on both a global and a single limited area model. The joint state method has been shown to outperform an analysis method that considers the limited area and the global model information separately. In the LETKF, only observations that fall into an area of fixed radius around a given grid point are allowed to affect the analysis at that location. The joint state vector at a given location is defined by concatenating all of the global and limited area model state information in the local region from which observations are considered. A background ensemble of joint state vectors is formed, and the LETKF is applied to produce an analysis ensemble of the joint state vectors, from which we select the analysis ensemble values only for the central grid point.

In our method, at grid points where c LAMs are defined, where $c > 1$, the joint-state analysis is performed once for each of the c Global Model - LAM pairs. These quantities are then combined into a composite **w**,**W** pair via a weighted average:

$$\mathbf{w}_n = \sum_{i=1}^{c} \alpha_i(\mathbf{x}_n)\mathbf{w}_{i,n} \qquad \text{Eq. 1}$$

Here the subscript n indexes grid point, and i indexes LAMs. The α-functions, which represent the weights of the average, are spatially dependent. Each LAM is most accurate at grid points near the center of its domain. As a result, when LAMs overlap, the weight **w,W** should not be treated as equals for the different LAMs, as some will have been derived using LAM data that is relatively closer to its domain center, and hence more accurate. To accommodate this consideration, the coefficients of our weighted average, denoted above by $\alpha_i(\mathbf{x}_n)$, are spatially dependent, and smoothly decrease from 1 to 0 as location moves towards the LAM boundaries. The same averaging procedure is also used to average the weight matrices, **W**.

For our idealized experiments, we use models introduced by Lorenz (2005), hereafter referred to as Lorenz Models II and III. These two models describe the behavior of a fictional atmospheric-like quantity on a circle of constant latitude using a system of autonomous ordinary differential equations. The main difference between the two models is that that Model III exhibits behavior on two spatial scales, while Model II has only one (Lorenz 2005). Model III governs our 'truth' dynamics, from which we create observations, and to which we compare our state estimates. Our global model, described by Lorenz model II, is run on a courser resolution grid. Each grid point of our global model coincides with every 4^{th} grid point of the true model. Our experiments test a system of two LAMS, both based on Model III, which are each defined on continuous subsets of the truth model grid (at the same resolution).

As a first test of the efficacy of our method, we compute the forecast improvements for both a regional and the global models over a given domain, when a second LAM is added to the system. In these experiments, each LAM covers 540 grid points, running from grid points 0 to 540, and 480 to 60 (a result of the periodic boundary conditions on a grid of 960 points). Both LAMS have the same model parameters as the perfect model. Observations are taken at regular intervals, once at every 64^{th} grid

point of the truth model. The analysis region extended for 40 truth model grid points in either direction from the location where the analysis is performed.

Figure 1 shows the effect of adding a second LAM on the analysis accuracy for the first LAM. Then the second LAM was added and a new analysis was produced by the procedure outlined above. As can be seen by comparing the red and the black curves, adding a second LAM in an adjoining region improves the analysis accuracy. We hypothesize that these improvements are due to the improved regional analyses in the overlapping regions, as well as to the improved global model analysis. The latter improvements are illustrated by Fig. 2.

The preliminary results suggest that the joint state method can be used to make numerical weather prediction more amenable to distributed parallel computing, as it would allow for the parallelization of not only the analysis, but also the model integration.

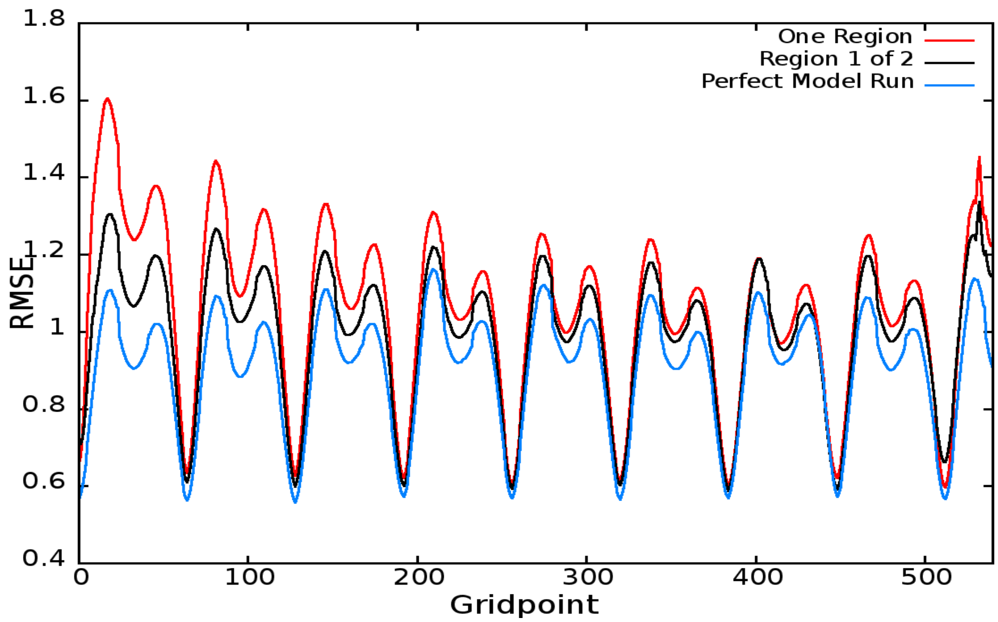

Figure 1. Temporally averaged root-mean-square analysis error in the LAM domain 0-540 for the perfect model (blue), for the LAM (red), and for the case when the second LAM is added (black).

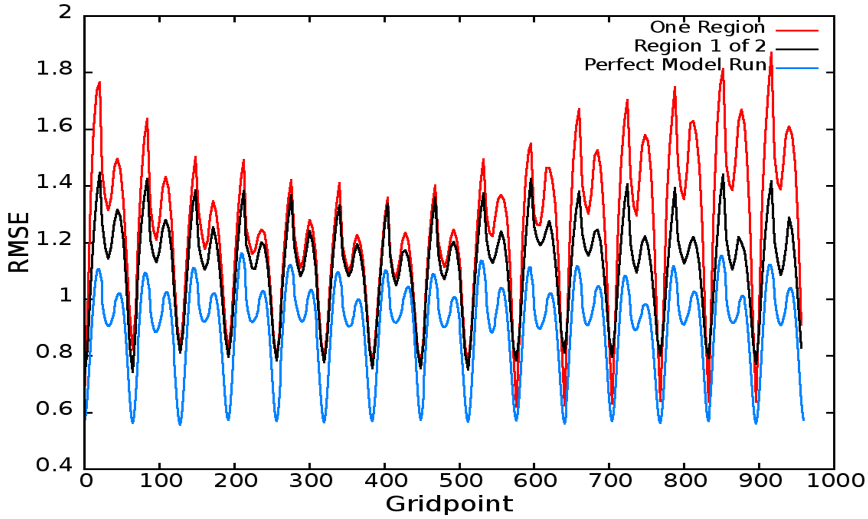

Figure 2. Same as Fig. 1, except for the global model domain.

Robust Ensemble Kalman Filter

The PI (Istvan Szunyogh) collaborated with a group of statisticians on the development of a robust ensemble Kalman filter for observation quality control. While robust statistics have an extensive mathematical literature, the application of those rigorous results to the different aspects of data assimilation is surprisiingly rare. The only exception has been observational quality control in the context of 4-dimensional variational data assimilation. The goal of our research has been to extend that theory to ensemble-based Kalman filter data assimilation schemes. The resulting theory was summarized in Roh et al. (2013) and has already been successfully tested with an implementation on our coupled global-limited-area data assimilation system.

While in a conventional data assimilation system observations that are significantly different from their predicted value are rejected, in the robust ensemble Kalman filter, the innovations associated with the suspect data are reduced based on a prescribed statistical model for the observation errors. In other words, the effect of the observations that strongly deviate from their predicted value on the state estimate is reduced, but not eliminated. The advantage of this approach is that instead of making a decision between assimilating the observation with a full weight or rejecting it based on a subjectively selected threshold value for the magnitude of the innovation, the weight of the observations is gradually reduced with increasing magnitude of the innovation. This property of the quality control is particularly advantageous for the assimilation of observations in the vicinity of a tropical cyclones, where the large magnitude of the innovations is not necessarily due to a gross error in the observations.

Results with the Coupled Global-Limited-Area Data Assimilation System

Experiments were conducted to evaluate the quality control algorithm for three types of tropical cyclone observations: TCVitals minimum sea level pressure (mSLP), QuikScat 10 m wind vectors, and reconnaissance dropsonde observations. In several single-time analysis experiments for Typhoon Sinlaku, the quality control algorithm was tested for each observation type individually and combined. The analysis results indicate that keeping the TCVital mSLP with low error (~0.5 hPa) and no quality control deepens the cyclone more than any other experiment, but special care must be taken when the

storm becomes too deep to be accurately resolved by the grid-spacing of the model. Keeping all direct TC observations results in the deepest simulated cyclone, but is not the best practice because of the inherent errors of the observations. Clipping the observations to a prescribed maximum value, as required by the robust ensemble Kalman filter, safeguards against blindly keeping truly erroneous observations. We have found that keeping the TCVitals unchanged and clipping the Quickscat and dropsonde observations provides the best analysis and forecast results. These results are illustrated by Figure 3, which shows the wind field and the sea level pressure for the single-time analyses. The Clipped and Kept experiments deepen the storm significantly over the Quality Control experiment and reposition the center of circulation to a position more consistant with the Best Track location.

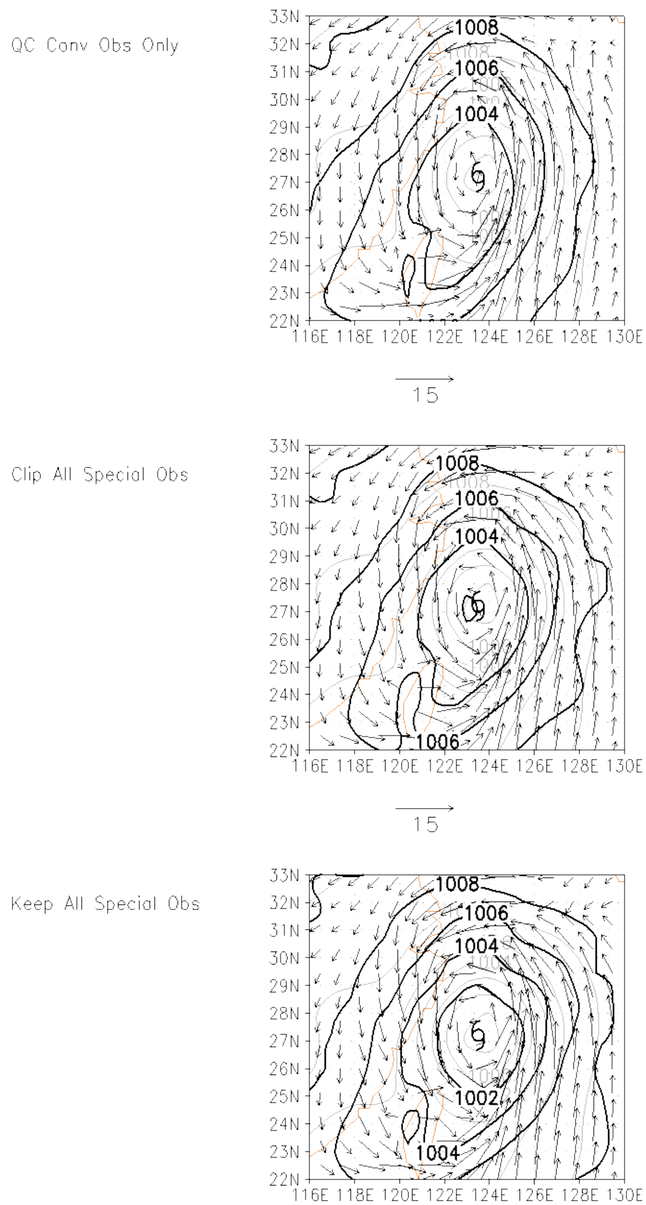

Figure 3. Mean sea level pressure contours for each experiment (black) and the 0.5° NCEP Operational Analysis (gray) with experiment wind vectors. The Best Track storm position is marked with a tropical cyclone symbol.

Deterministic forecasts from the single update experiments also indicate that the COMB 0.5 experiment (keeping TCV mSLP with 0.5 hPa error and clipping other special observations) improves the track forecast past day two (Fig. 4), while improving intensity forecasts in the first two days (Fig. 5). Clipping (CS) and Keeping (KS) all of the special observations resulted in similar forecasts, but the aforementioned reasons make those experiments less ideal than the combined quality control. The COMB 1.0 forecast (same as COMB 0.5, but with 1.0 hPa TCV error) does not significantly differ from the COMB 0.5 forecast for this particular time, but makes a significant difference in the analysis of intensity at other times (not shown for the single update experiment).

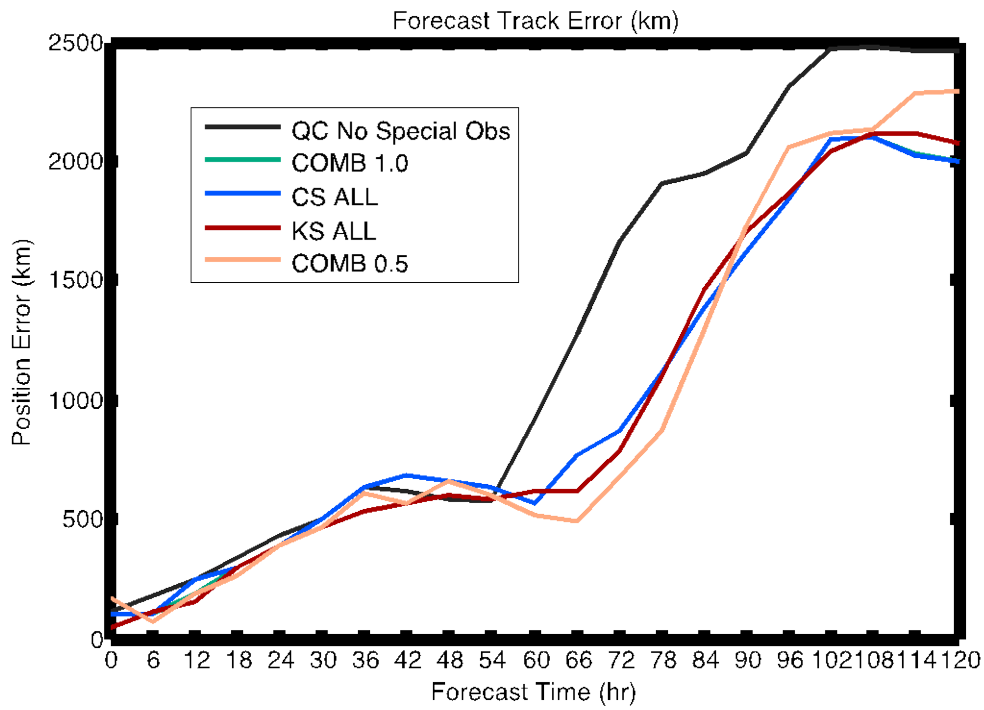

Figure 4. Forecast track error for single update experiments from 0000 UTC 16 Sep 2008.

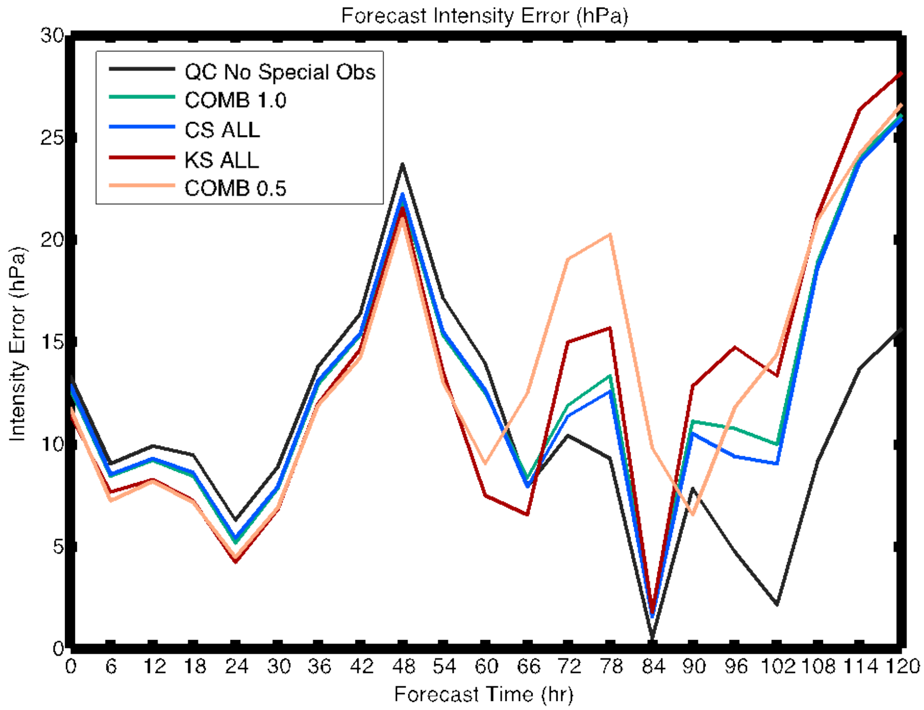

Figure 5. Forecast intensity error for single update experiments from 0000 UTC 16 Sep 2008.

Cycled experiments for Typhoon Sinlaku (2008) provide exceptional results with the Combined 0.5 quality control experiment (Fig. 6). When cycling the analysis, the combined experiment clips all observation innovations unless QuikSCAT observations are available in the vicinity of the cyclone, at which times the TCVital mSLP is kept with no quality control. Experiments including the quality controlled QuikSCAT observations with the kept TCVitals mSLP (not shown) indicate that the 10 m winds limit the depth of the simulated cyclone, supporting the decision to keep TCVitals observations when possible. The Combined 0.5 experiment captures the trend in intensity that none of the other experiment can, including the 0.5° NCEP Operational Analysis. The use of the Combined 0.5 quality control method improves the intensity analysis of Typhoon Sinlaku by as much as 40 hPa for a single update and 15 hPa on average over the traditional quality control methods with and without the special observations, as well as the global LETKF analysis. The average position error is also improved by the Combined 0.5 quality control experiment, as seen in Fig. 7 for all of the same experiments. Both of the Combined experiments improve the position of the analyzed storm over the NCEP Operational analysis and the traditional quality control experiments.

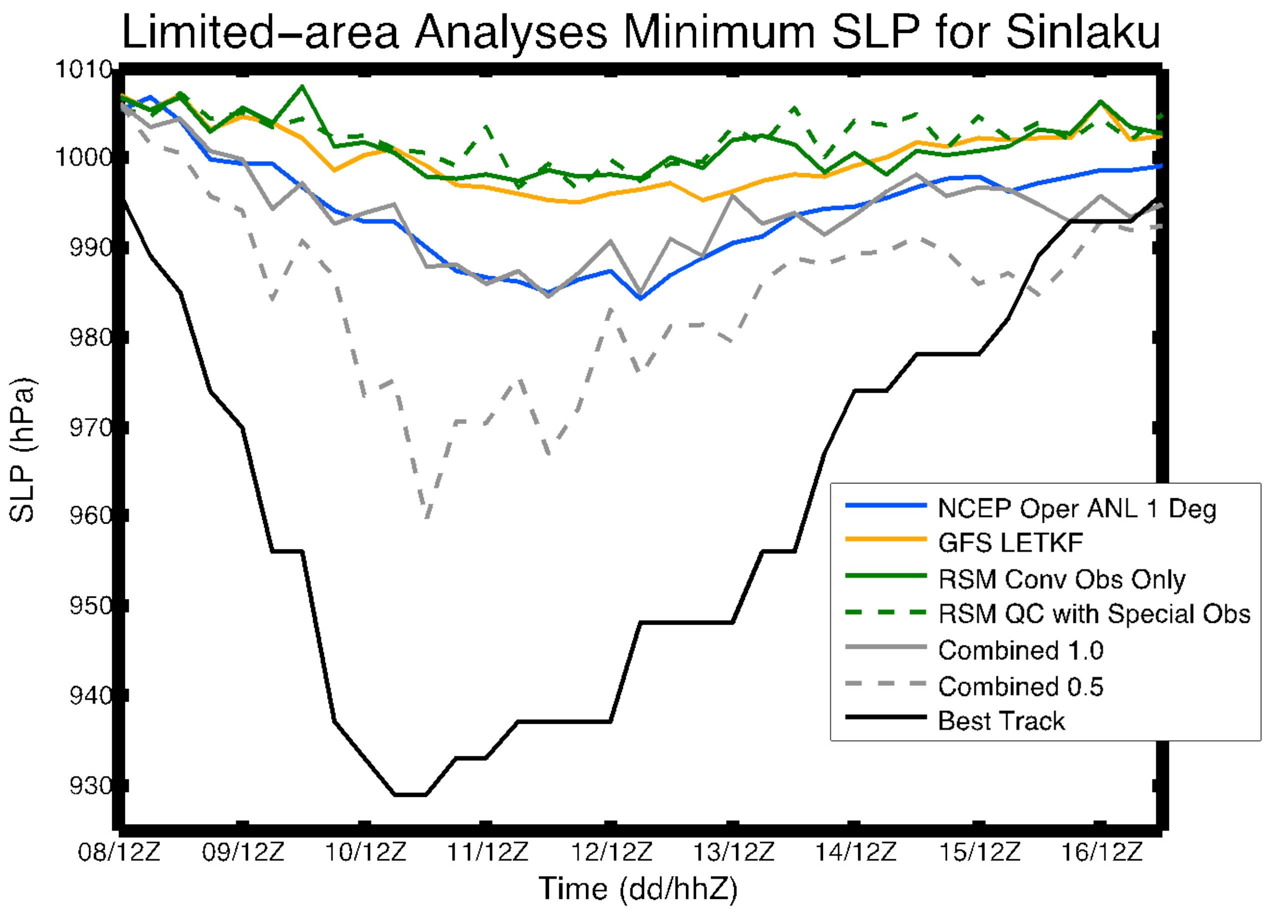

Figure 6. Cycled analysis intensity for Typhoon Sinlaku (2008).

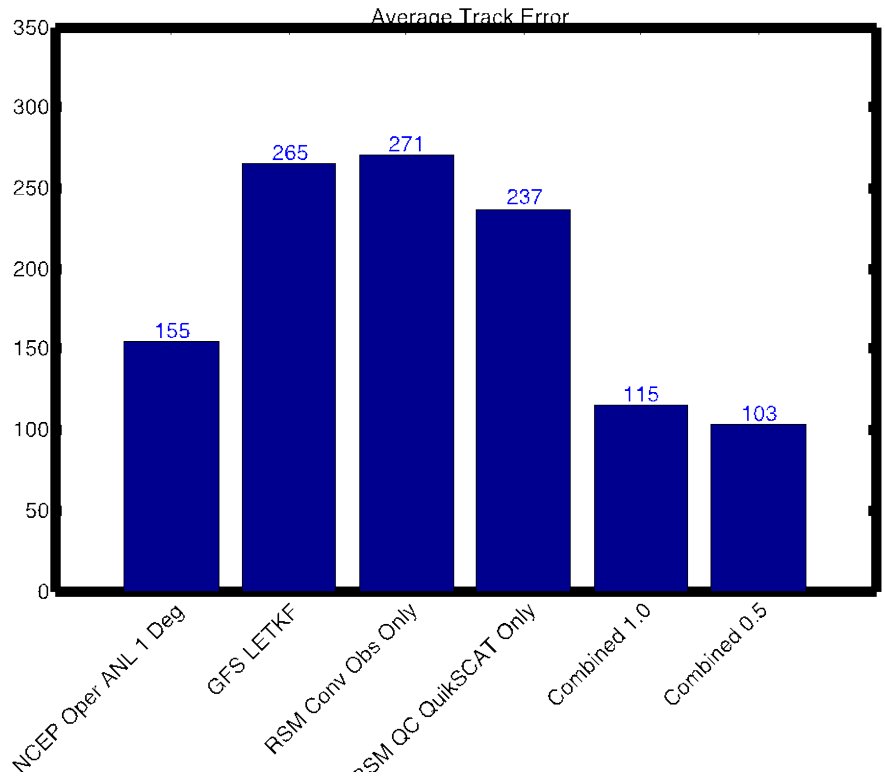

Figure 7. Track error (km) averaged over all cycles for Sinlaku.

IMPACT/APPLICATIONS

Our results suggests that using multiple limited area regions in the data assimilation process, which are available at no extra cost in the FNMOC forecast suite, can lead to improvements in both the limited area and the global analysis. Our results also indicate that the observation quality control algorithm plays a key role in the quality of the analysis of a tropical cyclone. Our plan for the next year is to focus on the implementation of our new techniques on the forecast system of the Navy for further testing.

REFERENCES

Holt, C. R., I. Szunyogh, and G. Gyarmati, 2012: Can a moderate resolution limited-area data assimilation system add value to the global analysis of tropical cyclones? *Mon. Wea. Rev,* **141**, 1866-1883.

Holt, C. R., 2011: *Testing a coupled global-limited area data assimilation system using observations from the 2004 Pacific typhoon season.* Master of Science Thesis, Texas A&M University, 60 pp.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112-126.

Lorenz, E. N., 2005: Designing Chaotic Models. *J. Atmos. Sci.,* **62,** 1574 – 1587.

Merkova, D., I. Szunyogh, and E. Ott, 2011: Strategies for coupling global and limited-area ensemble Kalman filter assimilation. *Nonlin. Processes Geophys.,* **18,** 415-430.

Roh, S., M. C. Genton, M. Jun, I. Szunyogh, and I. Hoteit, 2013: Observation quality control with a robust ensemble Kalman filter? *Mon. Wea. Rev,* **141**, (in press).

Yoon, Y.-N., B. R. Hunt, E. Ott, and I. Szunyogh, 2012: Ensemble regional data assimilation using joint states. *Tellus,* **64A**, 18407.